

文章编号: 1674-5566(2010)02-0157-05

奥利亚罗非鱼 MSTN 基因结构 及其 SNPs 的筛选

沈丽红¹, 熊良伟², 唐永凯³, 王帅兵²

(1. 浦东新区水产技术推广站, 上海 201200;

2. 江苏畜牧兽医职业技术学院, 江苏 泰州 225300;

3. 中国水产科学研究院淡水渔业研究中心, 江苏 无锡 214081)

摘要:通过 PCR 法从奥利亚罗非鱼 DNA 中扩增出 MSTN 基因。该序列全长 1 957 bp 含有 3 个外显子, 2 个内含子。外显子 I 379 bp 外显子 II 371 bp 部分外显子 III 145 bp 内含子 I 305 bp 内含子 II 751 bp 编码区 895 bp 共翻译 298 个氨基酸。通过随机测序法筛选 SNPs 获得了 6 个突变位点, 外显子 I、II 各有一个, 而内含子 II 则含有 4 个突变位点。外显子 I、II 的突变位点均导致翻译的氨基酸序列发生改变, 为错义突变。研究结果为 SNPs 位点与奥利亚罗非鱼生长性能关联分析奠定了基础。

关键词:奥利亚罗非鱼; MSTN 基因; SNPs

中图分类号: S 917 **文献标识码:** A

The gene structure and identification of single nucleotide polymorphisms of MSTN in *Oreochromis aureus*

SHEN Li-hong¹, XIONG Liang-wei², TANG Yong-kai³, WANG Shuai-bing²

(1. Shanghai Pudong New Area Aquatic Product Technical Advice Station, Shanghai 201200, China;

2. Jiangsu Animal Husbandry and Veterinary College, Taizhou 225300, China;

3. Freshwater Fisheries Research Center, Chinese Academy of Fishery Sciences, Wuxi 214081, China)

Abstract: *Oreochromis aureus* MSTN was isolated by using PCR technique. The whole sequence length was 1 957 bp including three exons and two introns. The sequence length of exon I and II was 379 bp, 371 bp respectively, and partial exon III was 145 bp. The sequence length of Intron I and II was 305 bp and 751 bp respectively. The ORF was 895 bp encoding 298 amino acids. Six SNPs was obtained through sequencing MSTN from 10 randomly selected individuals in *Oreochromis aureus*. Four SNPs were found in the Intron II, however, only one SNPs was observed in the Exon I and II respectively. The mutation sites in the Exon I and II changed for amino acids indicated that they were sense mutation. The SNPs might be useful for studying MSTN gene association to growth trait in *Oreochromis aureus*.

Key words: *Oreochromis aureus*; MSTN gene; SNPs

收稿日期: 2009-06-11

基金项目: 基本科研业务费专项资金 (2007JBFC03; 2007JBFC04); 国家科技基础条件平台 (2007DKA30470_003)

作者简介: 沈丽红 (1981-), 女, 助理工程师, 主要从事鱼类遗传育种方面的研究。E-mail: jianwz@sina.com

通讯作者: 唐永凯, E-mail: tangyk@ffrc.cn

肌细胞生长抑制素 (Myostatin, MSTN) 又称 GDF-8 (growth differentiation factor 8), 是近年来发现的一类重要的肌细胞生长负向调控因子, 属于转化生长因子超家族 (transforming growth factor- β superfamily, TGF β), 通过抑制分化生肌决定因子 (myogenic determination gene MyoD) 家族成员转录活性来控制肌细胞的生长发育, 它的表达量与肌肉重量的变化呈负相关^[1-2]。通过基因打靶技术敲除老鼠体内 MSTN 基因使其不能发挥作用, 结果发现突变型老鼠的体型显著大于野生型^[3-4]。在比利时蓝牛 (Belgian Blue) 和皮尔蒙特牛 (Piedmontese) 上, MSTN 的突变导致蛋白活性区的移码突变, 不能形成成熟的 MSTN 蛋白, 结果出现“双肌症状” (double muscle)^[5]。而在水产动物上, 目前已克隆出了虹鳟 (*Oncorhynchus mykiss*), 斑马鱼 (*Danio rerio*), 罗非鱼 (*Oreochromis mossambicus*), 金头鲷 (*Spanus aurata*), 鲈鱼 (*Dicentrarchus labrax*), 斑点鲰 (*Ictalurus punctatus*), 真鲷 (*Chrysophrys major*) 等鱼类的 MSTN^[6-11], 但有关 MSTN 基因结构研究则较少。

单核苷酸多态 (signal nucleotide polymorphisms SNPs) 是指在基因组水平上由单个核苷酸的变异所引起的一种 DNA 序列多态性, 是继限制性片段长度多态性 (restriction fragment length polymorphism, RFLP)、简单串联重复序列 (simple tandem repeat STR) 之后的第三代分子遗传标记, 在人类遗传性疾病的病因诊断、治疗和防治方面已受到广泛的重视和强有力的支持^[12]。在动物育种方面, SNPs 作为一种分子标记辅助育种强有力的工具也得到了广泛的应用, 将 SNPs 与经济性状进行相关性分析, 寻找影响性状的主效基因或与主效基因紧密连锁的候选基因, 将其运用于后续的育种实践中^[13-14]。在畜禽上, MSTN 的 SNPs 普遍存在, 而且有些位点和动物经济性状存在着显著性相关^[15-16], MSTN 突变的动物不仅骨骼肌肌群分布广泛, 而且肉质性状优良, MSTN 基因已成为肉质性状的主要标记基因^[17]。而在水产动物上, 将 MSTN 的 SNPs 与经济性状进行相关研究还未见报道。本研究旨在通过克隆罗非鱼的 MSTN, 分析其基因结构, 并寻找 SNPs 为 SNPs 位点与罗非鱼生长性能关联分析奠定了基础。

1 材料与方法

1.1 材料

1.1.1 实验鱼

罗非鱼取自中国水产科学研究院淡水渔业研究中心宜兴养殖基地。

1.1.2 试剂

限制性内切酶、pMD18-T 载体、胶回收试剂盒等购自宝生物工程 (大连) 有限公司; 质粒抽提试剂盒、Taq 酶购自上海捷瑞生物工程有限公司; 大肠杆菌 JM109 为本实验室保存。

1.1.3 仪器:

PCR 仪为 Eppendorf Mastercycler personal

1.1.4 引物:

用于实验的所有引物见表 1, 碱基位置以序列 5' 端第一个碱基为“1”计算。根据莫桑比克罗非鱼 MSTN 的 cDNA (GenBank: AF197193) 以及其它鱼类 MSTN 的 cDNA 和 DNA 序列设计引物, 分段扩增罗非鱼 MSTN 序列, 共 3 个外显子, 两个内含子。其中 Forward primer¹ (F1), Reverse primer¹ (R1) 扩增 Exon I - Intron I 片段, Forward primer² (F2), Reverse primer² (R2) 扩增 Exon II 片段, Forward primer³ (F3), Reverse primer³ (R3) 扩增 Intron II 片段。所有引物均由上海捷瑞生物工程有限公司合成。

表 1 实验中使用的引物及相应的扩增产物
Tab. 1 Primers and their amplified production in the study

引物	序列	位置	产物
F1	GAGACAATGCATCTGTCTCAGATCG	1-25	Exon I - Intron I
R1	GCGACTGGCTTGAAACTTCTGC	746-767	767bp
F2	ATTTTCTTTTCCCATGTCTCCG	666-687	Exon II
R2	AAACGATGAACAATAGGTGGAGC	1120-1142	477bp
F3	CACATACGAATCCGCTCCTTGA	870-891	Intron II
R3	GAGGACTTGGCTGGACTGGA	1936-1957	1088bp

1.2 方法

1.2.1 基因组 DNA 的提取

从鱼尾静脉抽血 0.2~0.5 mL, 用经典的苯酚氯仿法抽提 DNA, 并溶解于 TE 中, 电泳检测确定 DNA 质量, 在紫外分光光度计 (Eppendorf) 上测定 DNA 浓度, DNA 使用液浓度为 50 ng/ μ L。

1.2.2 奥利亚罗非鱼 MSTN 基因的分离

根据设计的引物分别扩增出奥利亚罗非鱼 MSTN 内含子、外显子序列。反应条件为: 94 °C 3 min, 然后 30 次循环 94 °C 1 min, 55 ~ 58 °C 1 min, 72 °C 1 ~ 2 min, 最后 72 °C 10 min, 4 °C 保存。扩增液用 1.0% 的琼脂糖凝胶电泳分离, 切胶回收, 连接转化, 蓝白斑挑选阳性克隆, 酶切鉴定, 送南京博亚生物有限公司测序。

1.2.3 多态位点的筛选

随机选取 10 尾奥利亚罗非鱼, 以 DNA 为模板, 用上述扩增 MSTN 内外显子的 3 对引物扩增, 经克隆、测序比对, 寻找 SNPs 位点。

1.2.4 序列分析和数据处理

序列拼接以及分析使用 DnaStar 软件, 序列比对使用 ClustaW 软件。

2 结果

2.1 奥利亚罗非鱼 MSTN 基因的分离和分析

使用引物 F1-R1、F2-R2、F3-R3 扩增奥利亚

罗非鱼 Exon I -Intron I、Exon II、Intron II, 获得 DNA 序列片段分别为 767 bp、477 bp 和 1 088 bp, 经过拼接获得奥利亚罗非鱼 MSTN 基因总共 1 957 bp (GeneBank No. FJ972684), 含有 3 个外显子和 2 个内含子。Exon I 379 bp, Intron I 305 bp, Exon II 371 bp, Intron II 751 bp, 部分 Exon III 145 bp, 阅读框长 895 bp, 编码 298 个氨基酸, 内含子均符合 GT-AG 原则 (图 1)。与吉富罗非鱼 (GeneBank No. FJ972683)、莫桑比克罗非鱼 MSTN (GeneBank No. AH006117) 序列比对显示, 阅读框核苷酸序列长度一致, 相似性分别为 99.3%、98.9%, 编码的氨基酸序列相似性分别为 99%、97.3%。

2.2 多态位点的筛选

使用不同引物对, 分别在随机选取的 10 尾奥利亚罗非鱼基因组中扩增相应片段, 克隆测序后, 使用 ClustaW 进行比对, 获得了 6 个突变位点 (图 1)。外显子 I、II 各有一个, 内含子 II 含有 4 个突变位点, 而且外显子 I、II 的 SNPs 均导致编码的氨基酸序列发生变异, 为错义突变。

```

gagacaATG CAT CTG TCT CAG ATC GTG CTG TAT CTG AGC TTG CTG ACT GCG TTG GGT CCA 60
  I      M H L S Q I V L Y L S L L T A L G P 20
61 GTA GTT CTG AGT GAC CAA GAG GCG CAC CAG CAG CCT TCC GTC AGC ACC CCA GTA GAC ACG 120
21 V V L S D Q E A H Q Q P S V S T P V D T 40
121 GAT CAA TGC GCT ACC TGC GAG GTC CGG CAG CAG ATT AAA ACC ATG CGA TTA AAC GCG ATA 180
41 D Q C A T C E V R Q Q I K T M R L N A I 60
181 AAG TCT CAG ATT CTG AGC AAG CTG CGA ATG AAG GAA GCC CCC AAC ATT AGC AGA GAG ATC 240
61 K S Q I L S K L R M K E A P N I S R E I 80
241 GTG AAG CAG CTC CTG CCC AAA GCG CCG CCG CTG CAG CAG CTT CTG GAT CAG TAC GAC GTG 300
81 V K Q L L P K A P P L Q Q L L D Q Y D V 100
301 CTG GGA GAT GAC AAC AGG GAA GAA GTT CTG GAG GAC GAC GAC GAG CAC GCA ACC ACG GAG 360
101 L G D D N R E E V L E D D D E H A T T E 120
361 ACA ATT GTA ATG ATG GCA ACT GAA Cgtagtgtgaattttcttattttttgtctctcgtgagcgcactcgaa
cgtctccaaataagtttttacgcacgtgtagagcgcacaggagagttatttttagagagcactcagttttccaccaaggt
tttgtttaaagtgtcatgtaaaagattcaaacactttatgagacatggtgtttgggcagagagtgcgccgagtgattacgca
tgagttttacgcaatcatgcgcaaacgcgccactttatttctcagcgcgatcatggaattaattttctttccatgtctccgag
CT GAT TCC GCT GTC CAG GTG GAC GGG CAA CCA AAG 420
121 T I V M M A T E P D S A V Q V D G Q P K 140
421 TGC TGC TTT TTC TCA ATT ACG CAG AAG TTT CAA GCC AGT CGC GTA GTT CGA GCG CAG CTT 480
141 C C F F S I T Q K F Q A S R V V R A Q L 160
481 TGG GTG CAT CTG CGT CCA TCG GAA GAA GTG ACC ACC GTG TTC CTG CAA ATC TCC CGG CTT 540
161 W V H L R P S E E V T T V F L Q I S R L 180

```

```

541 ATA CCG GTC ACA GAC GGG AAC AGA CAC ATA CGA ATC CGC TCC TTG AAG ATC GAC GTG AAT 600
181 I P V T D G N R H I R I R S L K I D V N 200
601 GCC GGG GCC AGC TCT TGG CAA AGT ATA GAC GTC AAG CAA GTG TTG ACT GTG TGG CTG CGG 660
201 A G A S S W Q S I D V K Q V L T V W L R 220
661 CAG CCG GAG ACC AAC TGG GGC ATC GAG ATC AAC GCT TTC GAT TCG AGG GGA AAT GAC TTA 720
221 Q P E T N W G I E I N A F D S R G N D L 240
721 GCT GTG ACC TTC GCA GAG CCG GGA GAG GGT CTG gtgagctcaactttactgaaatttaacagggga
gcttgggtaagtttcacgttgaaacgatgaacaataggtggagcttggctcatatgtgcagagcagtatttgcgcttctg
tcacaatcgggtgcatatctagccatttcaaagcgcctaagttcagaaaactgtaatgcgtctctgtacagaggggtgagact
ttaacttacagtgtgcagagaggcttgcaggctcagaggcgtttaaacattctgcatttactgctggaagaaggaattg
ggatataaaagacttttgaagctgtaatacacacagactccatgtaggcgcacacctgcatttggacaagcggccggctc
acaataacctcaaatcttattgcttgattacataaaagttcacctgctccacattcctttaaagtattcagtggtgctgcaaatca
gagtaattgccggcacatacacacacagagctactatcaatgacaacatacatcacagagctctaataatggcaccggc
actggggcaattagaataactagacagctgtggctaaccaaccatcagatgatttaatttctgataactcaaagtttcataata
ttatcgtccatgtttgaaaactgactgccgttaggggcagcggattaaaagatgaaaagctgaatgctgtcaccgcctcctc
ctgcaaaagctgttattttcaaatgtttcacacacactctctgttctag CAA CCG TTC ATG GAG GTG AAG ATT 780
241 A V T F A E P G E E G L Q P F M E V K I 260
781 TCA GAG GGC CCC AGG CGT GCC CGG AGA GAC TCG GGC CTG GAC TGC GAT GAG AAC TCT CCA 840
261 S E G P R R A R R D S G L D C D E N S P 280
841 GAG TCC CGG TGC TGT CGC TAC CCA CTC ACT GTG GAC TTC GAG GAC TTT GGC TGG GAC TGG 900
281 E S R C C R Y P L T V D F E D F G W D W 300
901 A 901

```

图 1 奥利亚罗非鱼 MSTN 基因序列

Fig 1 Genomic sequence of MSTN gene in *Oreochromis aureus*

注: 5'调控区及内含子用小写字母表示, 外显子用大写字母表示, 内含子与外显子交接处用黑体和斜体表示, SNPs位点以方框和黑体表示。

3 讨论

首次分离了奥利亚罗非鱼 MSTN 基因, 结果表明该基因具有 3 个外显子, 2 个内含子, 基因结构与虹鳟^[6]、斑马鱼^[7]、真鲷^[11]等一致。与吉富罗非鱼、莫桑比克罗非鱼 MSTN 序列比较发现, 编码区核苷酸序列相似性高达 99.3%、98.9%, 这表明 MSTN 在种间高度保守, 这也与叶寒青等^[11]在真鲷中的分析结果相一致。根据这一特性, 我们设计的 3 对引物可以直接从 DNA 中扩增出 MSTN 的编码区, 而无需从 RNA 中来扩增。对于同源性较高的基因, 完全可以利用同源克隆的策略来获得目的基因, 既节省了研究时间又节省了研究经费。

在 SNPs 的查找中, 通过测序比较得到了 6 个 SNPs 但还需要在群体中进一步验证, 因为通

过随机测序获得的 SNPs 可能是假阳性, 在 PCR 扩增, 克隆, 测序中也可导致单碱基突变。在 SNPs 的发掘中, 单碱基突变出现的概率应大于 1/10 (即 10 个体中有 1 个以上个体出现单碱基突变) 才有可能为 SNPs 位点。然而筛选到的 6 个 SNPs 其中 4 个位于内含子内, 可能由于内含子不参与功能基因的编码, 所受的选择压力小, 因而比外显子更容易积累较大的变异, 这与于凌云^[18]等在大口黑鲈 MyoD 上的研究结果有些相似, 只在 MyoD 的内含子上找到了 7 个 SNPs, 另外 2 个 SNPs 位于外显子上, 而且均导致编码的氨基酸发生变异, 至于能否引起 MSTN 功能的变异还有待进一步的验证, 以便更好地应用于后续分子标记辅助育种上, 确定成熟的分子标记辅助育种方案。

参考文献:

- [1] Oldham J M, Martyn J A K, Shama M, et al. Molecular expression of myostatin and MyoD is greater in double-muscled than normal-muscled cattle fetuses [J]. *Am J Physiol Regul Integr Comp Physiol* 2001, 280: 1488—1493.
- [2] Langley B, Thomas M, Bishop A, et al. Myostatin inhibits myoblast differentiation by down-regulating MyoD expression [J]. *J Biol Chem* 2002, 277(51): 49831—49840.
- [3] Szabo G, Dallmann G, Muller G, et al. A deletion in the myostatin gene causes the compact (Cmpt) hypemuscular mutation in mice [J]. *Mamm Genome* 1998, 9: 671—672.
- [4] McPherron A C, Lawler A M, Lee S J. Regulation of skeletal muscle mass in mice by a new TGF β superfamily member [J]. *Nature* 1997, 387: 83—90.
- [5] Kambadur R, Shama M, Smith T P, et al. Mutations in myostatin (GDF-8) in double-muscled Belgian Blue and Piedmontese cattle [J]. *Genome Res* 1997, 7: 910—916.
- [6] Dilip K G, Scott A G, Buel D R. Identification, Characterization and Quantitative Expression Analysis of Rainbow Trout Myostatin^{-1a} and ^{-1b} Genes [J]. *Journal of Endocrinology* 2006, 190: 879—888.
- [7] Xu C, Wu G, Zohar Y, et al. Analysis of myostatin gene structure, expression and function in zebrafish [J]. *J Exp Biol* 2003, 206: 4067—4079.
- [8] Rodgers B D, Weber G M, Sullivan C V, et al. Isolation and characterization of myostatin complementary deoxyribonucleic acid clones from two commercially important fish: *Oreochromis mossambicus* and *Morone chrysops* [J]. *Endocrinology* 2001, 142(4): 1412—1428.
- [9] Weber T E, Small B C, Bosworth B G. Lipopolysaccharide regulates myostatin and MyoD independently of an increase in plasma cortisol in channel catfish (*Ictalurus punctatus*) [J]. *Domest Anim Endocrinol* 2005, 28(1): 64—73.
- [10] Tenova G, Bernardini G, Binelli G, et al. cDNA encoding sequences for myostatin and FGF6 in sea bass (*Dicentrarchus labrax* L.) and the effect of fasting and refeeding on their abundance levels [J]. *Domest Anim Endocrinol* 2006, 30(4): 304—319.
- [11] 叶寒青, 陈松林, 真鲷肌肉生长抑素 MSTN 基因的克隆及表达分析 [J]. *高技术通讯*, 2006, 16(7): 718—724.
- [12] Wang D Q, Fan J B, Siao C J, et al. Large-scale identification mapping and genotyping of single nucleotide polymorphisms in the human genome [J]. *Science* 1998, 280: 1077—1082.
- [13] 王宝维, 魏笑笑, 贾晓晖, 等. 鲁西斗鸡 IGF-1 基因多态性与体型指标关系的研究 [J]. *中国家禽*, 2008, 30(12): 20—24.
- [14] 黄丹丽, 陈仁金, 杨章平, 等. 4 个绵羊品种 Leptin 基因部分片段的 SNPs 多态性与生长发育性状的相关性分析 [J]. *畜牧兽医学报*, 2008, 39(12): 1640—1646.
- [15] 顾志良, 张海峰, 朱大海, 等. 鸡 myostatin 基因单核苷酸多态性的群体遗传学分析 [J]. *遗传学报*, 2002, 29(7): 599—606.
- [16] 郭丹, 曹阳, 张嘉保, 等. 草原红牛 MSTN 基因第一外显子 SNPs 多态性分析 [J]. *现代畜牧兽医*, 2007, 11: 12—14.
- [17] 胡兰, 郭东新, 胡锐, 等. 大骨鸡中 MSTN 基因表达规律性的研究 [J]. *动物科学与动物医学*, 2003, 20(11): 42—44.
- [18] 于凌云, 白俊杰, 叶星, 等. 大口黑鲈 MyoD 基因结构和单核苷酸多态性位点的筛选 [J]. *水产学报*, 2009, 33(1): 1—8.